

INVITED COMMENTARY

Is Bigger Actually Better? The Strengths and Limitations of Large Registry Data

Christopher P. Twine ^{a,*}, Regula S. von Allmen ^b

^a North Bristol NHS Trust, Bristol, UK

^b Clinic for Vascular Surgery, Kantonsspital St Gallen, St Gallen, Switzerland

As publications using data from huge registries become more popular, so does the belief that the strength of recommendation that can be made from the results is greater than from smaller cohorts. This very large dataset from Hamburg examines gender differences in peripheral endovascular revascularisation for peripheral arterial disease.¹ Such a large dataset is undoubtedly important, but potential sources of error and a low number of events for some significant results means that this study has good examples of the limits inherent to larger retrospective cohorts.

There have been several smaller retrospective cohort studies and secondary analyses of randomised trials showing differences in patterns of disease, treatment, and outcomes between male and female patients.² Although larger datasets with huge sample sizes appear to reduce sampling and selection error, it is a fact that they will inevitably be associated with the systematic exclusion of some relevant information.³

For example, in this paper pseudoaneurysms were more common in females than males, but there are no data on the type of intervention, sheath size, or closure device use in the female and male cohorts. The results cannot be fully adjusted in the way in which a smaller series with more detailed data fields could be, so even though such effects may have been mitigated by the larger number of patients we can never be sure. A propensity scored analysis of a complete dataset of 24,000 patients would be more powerful in further mitigating bias and confounding.

Another example to illustrate that even big data might be associated with a potential selection error is the fact that only inpatient procedures were available in this study, which is relevant when looking at gender differences. Isn't it possible that patients' sex might have influenced the decision to go for an either inpatient or outpatient procedure?

Most importantly, the large number of patients can also inflate the importance of (very) small effects. Again looking

at the example of pseudoaneurysms, around 330 female patients would need to be treated to develop one more pseudoaneurysm than a male patient. This result is only significant because of the size of the dataset, but how important is it in the real world? Most people would not expect to change their practice based on a number needed to treat of 330. The other conclusion drawn is that female patients are more at risk of social isolation; however, this trend was observed in 0.3% of women vs. 0% of men. The importance of this for a single vascular unit is questionable.

Where a large dataset like this excels is in the reduction of statistical test error. Therefore, the trends observed and refuted here are highly likely to be real. An example would be that female patients did not have a higher 30 day mortality in this paper, which contradicts previous smaller datasets.⁴ This now makes the previously published results appear confounded. So this dataset can draw strong conclusions about the patterns of clinical presentation and outcomes of female patients for endovascular intervention, and adds some definitive knowledge over the large number of smaller datasets already published.

REFERENCES

- Behrendt CA, Bischoff MS, Schwaneberg T, Böckler D, Hohnhold R, Diener H, et al. Population-based analysis of gender disparities in 23,715 percutaneous endovascular revascularisations in the metropolitan area of Hamburg. *Eur J Vasc Endovasc Surg* 2019;57: 658–65.
- Jelani QU, Petrov M, Martinez SC, Holmvang L, Al-Shaibi K, Alasnag M. Peripheral arterial disease in women: an overview of risk factor profile, clinical features, and outcomes. *Curr Atheroscler Rep* 2018;20:40.
- Kaplan RM, Chambers DA, Glasgow RE. Big data and large sample size: a cautionary note on the potential for bias. *Clin Transl Sci* 2014;7:342–6.
- Hultgren R, Olofsson P, Wahlberg E. Sex-related differences in outcome after vascular interventions for lower limb ischemia. *J Vasc Surg* 2002;35:510–6.

DOI of original article: <https://doi.org/10.1016/j.ejvs.2018.10.021>

* Corresponding author. North Bristol NHS Trust, Bristol, UK.

E-mail address: chris_twine@hotmail.com (Christopher P. Twine).

1078-5884/© 2018 European Society for Vascular Surgery. Published by Elsevier B.V. All rights reserved.

<https://doi.org/10.1016/j.ejvs.2018.11.006>